



## SEQUENCE LISTING

<110> Protalix Ltd.  
Shaaltiel, Yoseph  
Baum, Gideon  
Hashmueli, Sharon  
Lewkowicz, Ayala  
Bartfeld, Daniel

<120> PRODUCTION OF HIGH MANNOSE PROTEINS IN PLANT CULTURE

<130> 30570

<150> IL 155588

<151> 2003-04-27

<150> PCT/IL2004/000181

<151> 2004-02-24

<160> 14

<170> PatentIn version 3.5

<210> 1

<211> 22

<212> PRT

<213> Artificial sequence

<220>

<223> ER signal peptide

<400> 1

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser  
1 5 10 15

Leu Ser Ser Ala Glu Phe  
20

<210> 2

<211> 7

<212> PRT

<213> Artificial sequence

<220>

<223> Vacuolar targeting signal from Tobacco chitinase A

<400> 2

Asp Leu Leu Val Asp Thr Met  
1 5

<210> 3

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Single strand DNA oligonucleotide

<400> 3

cagaattcgc ccgcccctgc a

21

<210> 4

<211> 22

<212> DNA

<213> Artificial sequence

<220>  
 <223> Single strand DNA oligonucleotide  
 <400> 4  
 ctcagatctt ggcgatgcca ca 22

<210> 5  
 <211> 19  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> Single strand DNA oligonucleotide  
 <400> 5  
 ctcagaagac cagagggct 19

<210> 6  
 <211> 17  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> Single strand DNA oligonucleotide  
 <400> 6  
 caaagcggcc atcgtgc 17

<210> 7  
 <211> 1491  
 <212> DNA  
 <213> Homo sapiens  
 <400> 7  
 gcccgcccct gcatccctaa aagcttcggc tacagctcgg tgggtgtgtg ctgcaatgcc 60  
 acatactgtg actcctttga cccccgacc ttctctgccc ttggtacctt cagccgctat 120  
 gagagtacac gcagtgggag acggatggag ctgagtatgg ggcccatcca ggctaatacac 180  
 acgggacacag gcctgctact gacctgcag ccagaacaga agttccagaa agtgaaggga 240  
 ttggaggagg ccatgacaga tgctgctgct ctcaacatcc ttgccctgtc accccctgcc 300  
 caaaatttgc tacttaaate gtacttctct gaagaaggaa tcggatataa catcatccgg 360  
 gtacccatgg ccagctgtga cttctccatc cgcacctaca cctatgcaga caccctgat 420  
 gatttccagt tgcacaactt cagcctccca gaggaagata ccaagctcaa gataccctg 480  
 attcaccgag ccctgcagtt ggcccagcgt ccggtttcac tccttgccag ccctggaca 540  
 tcaccactt ggctcaagac caatggagcg gtgaatggga aggggtcact caaggacag 600  
 cccggagaca tctaccacca gacctgggcc agatactttg tgaagttcct ggatgcctat 660  
 gctgagcaca agttacagtt ctgggcagtg acagctgaaa atgagccttc tgctgggctg 720  
 ttgagtggat accccttcca gtgcctgggc ttcaccctg aacatcagcg agacttcatt 780  
 gcccgtagc taggtcctac cctcgccaac agtactcacc acaatgtccg cctactcatg 840  
 ctggatgacc aacgcttgct gctgccccac tgggcaaagg tggtagtgac agaccagaa 900  
 gcagctaaat atgttcatg cattgtgtga cattggtacc tggactttct ggctccagcc 960  
 aaagccaccc taggggagac acaccgctg ttccccaaca ccatgctctt tgctcagag 1020

```

gcctgtgtgg gctccaagtt ctgggagcag agtgtgcggc taggctcctg ggatcgaggg 1080
atgcagtaca gccacagcat catcacgaac ctctgtacc atgtggtcgg ctggaccgac 1140
tggaaccttg ccctgaaccc cgaaggagga cccaattggg tgcgtaactt tgtcgacagt 1200
cccatcattg tagacatcac caaggacacg ttttaciaaac agcccatgtt ctaccacctt 1260
ggccacttca gcaagttcat tcctgagggc tcccagagag tggggctggt tgccagtcag 1320
aagaacgacc tggacgcagt ggcaactgat catcccgatg gctctgctgt tgtggtcgtg 1380
ctaaaccgct cctctaagga tgtgcctctt accatcaagg atcctgctgt gggcttcctg 1440
gagacaatct cacctggcta ctccattcac acctacctgt ggcatcgcca g 1491

```

```

<210> 8
<211> 496
<212> PRT
<213> Homo sapiens

```

```

<400> 8

```

```

Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser Val Val Cys
1           5           10           15

```

```

Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro Thr Phe Pro
          20           25           30

```

```

Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser Gly Arg Arg
          35           40           45

```

```

Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr Gly Thr Gly
50           55           60

```

```

Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys Val Lys Gly
65           70           75           80

```

```

Phe Gly Gly Ala Met Thr Asp Ala Ala Ala Leu Asn Ile Leu Ala Leu
          85           90           95

```

```

Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe Ser Glu Glu
100           105           110

```

```

Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro Met Ala Ser Cys Asp Phe
115           120           125

```

```

Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr Pro Asp Asp Phe Gln Leu
130           135           140

```

```

His Asn Phe Ser Leu Pro Glu Glu Asp Thr Lys Leu Lys Ile Pro Leu
145           150           155           160

```

```

Ile His Arg Ala Leu Gln Leu Ala Gln Arg Pro Val Ser Leu Leu Ala
165           170           175

```

```

Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys Thr Asn Gly Ala Val Asn
180           185           190

```

Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly Asp Ile Tyr His Gln Thr  
 195 200 205  
 Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp Ala Tyr Ala Glu His Lys  
 210 215 220  
 Leu Gln Phe Trp Ala Val Thr Ala Glu Asn Glu Pro Ser Ala Gly Leu  
 225 230 235 240  
 Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly Phe Thr Pro Glu His Gln  
 245 250 255  
 Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro Thr Leu Ala Asn Ser Thr  
 260 265 270  
 His His Asn Val Arg Leu Leu Met Leu Asp Asp Gly Arg Leu Leu Leu  
 275 280 285  
 His Trp Ala Lys Val Val Leu Thr Asp Pro Glu Ala Ala Lys Tyr Val  
 290 295 300  
 His Gly Ile Ala Val His Trp Tyr Leu Asp Phe Leu Ala Pro Ala Lys  
 305 310 315 320  
 Ala Thr Leu Gly Glu Thr His Arg Leu Phe Pro Asn Thr Met Leu Phe  
 325 330 335  
 Ala Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu Gln Ser Val Arg  
 340 345 350  
 Leu Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His Ser Ile Ile Thr  
 355 360 365  
 Asn Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp Asn Leu Ala Leu  
 370 375 380  
 Asn Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe Val Asp Ser Pro  
 385 390 395 400  
 Ile Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys Gln Pro Met Phe  
 405 410 415  
 Tyr His Leu Gly His Phe Ser Lys Phe Ile Pro Glu Gly Ser Gln Arg  
 420 425 430  
 Val Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp Ala Val Ala Leu  
 435 440 445  
 Met His Pro Asp Gly Ser Ala Val Val Val Val Leu Asn Arg Ser Ser  
 450 455 460  
 Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val Gly Phe Leu Glu

465                      470                      475                      480

Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu Trp His Arg Gln  
                                  485                      490                      495

<210> 9  
 <211> 338  
 <212> DNA  
 <213> Cauliflower mosaic virus

<400> 9  
 ttttcacaaa gggtaatatc gggaaacctc ctcggattcc attgcccagc tatctgtcac 60  
 ttcacgaaa ggacagtaga aaaggaaggt ggctcctaca aatgccatca ttgcgataaa 120  
 ggaaaggcta tcgttcaaga tgcctctacc gacagtggtc ccaaagatgg acccccaccc 180  
 acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaagca agtggattga 240  
 tgtgatattc cacttgacgt aagggatgac gcacaatccc actatccttc gcaagaccct 300  
 tcctctatat aaggaagttc atttcatttg gagaggac 338

<210> 10  
 <211> 66  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Nucleic acid sequence encoding the ER signal peptide

<400> 10  
 atgaagacta atctttttct ctttctcatc ttttcacttc tcctatcatt atcctcggcc 60  
 gaattc 66

<210> 11  
 <211> 21  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Nucleic acid sequence encoding the vacuolar targeting sequence

<400> 11  
 gatcttttag tcgatactat g 21

<210> 12  
 <211> 167  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Nucleic acid sequence of the Agrobacterium tumefaciens terminator

<220>  
 <221> misc\_feature  
 <222> (162)..(162)  
 <223> n is a, c, g, or t

<400> 12  
 taatttcattg atctgttttg ttgtattccc ttgcaatgca gggcctaggg ctatgaataa 60  
 agttaatgtg tgaatgtgtg aatgtgtgat tgtgacctga agggatcacg actataatcg 120

tttataataa acaaagactt tgtcccaaaa accccccccc cngcaga 167

<210> 13  
 <211> 2186  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> nucleic acid sequence encoding high mannose human  
 glucocerebrosidase (GCD)

<220>  
 <221> misc\_feature  
 <222> (2181)..(2181)  
 <223> n is a, c, g, or t

<400> 13  
 ttttcacaaa gggtaatatc gggaaacctc ctcggattcc attgcccagc tatctgtcac 60  
 ttcacgaaa ggacagtaga aaaggaaggt ggctcctaca aatgccatca ttgcgataaa 120  
 ggaaaggcta tcgttcaaga tgctcttacc gacagtggtc ccaaagatgg acccccaccc 180  
 acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaagca agtggattga 240  
 tgtgatattc cactgacgt aagggatgac gcacaatccc actatccttc gcaagaccct 300  
 tcctctatat aaggaagttc atttcatttg gagaggacag gcttcttgag atccttcaac 360  
 aattaccaac aacaacaaac aacaaacaac attacaatta ctatttaca ttacagtcga 420  
 gggatccaag gagatataac aatgaagact aatctttttc tctttctcat cttttcactt 480  
 ctctatcat tatcctcggc cgaattcgcc cgccctgca tccctaaaag cttcggctac 540  
 agctcgggtg tgtgtgtctg caatgccaca tactgtgact cctttgacct cccgacctt 600  
 cctgcccttg gtaccttcag ccgctatgag agtacacgca gtgggacgac gatggagctg 660  
 agtatggggc ccatccaggc taatcacacg ggcacaggcc tgctactgac cctgcagcca 720  
 gaacagaagt tccagaaagt gaagggattt ggaggggcca tgacagatgc tgctgctctc 780  
 aacatccttg ccctgtcacc ccctgcccac aatttgctac ttaaatacgta cttctctgaa 840  
 gaaggaatcg gatataacat catccgggta cccatggcca gctgtgactt ctccatccgc 900  
 acctacacct atgcagacac ccctgatgat ttccagttgc acaacttcag cctcccagag 960  
 gaagatacca agctcaagat acccctgatt caccgagccc tgcagttggc ccagcgtccc 1020  
 gtttactcc ttgccagccc ctggacatca cccacttggc tcaagaccaa tggagcggtg 1080  
 aatgggaagg ggtcactcaa gggacagccc ggagacatct accaccagac ctgggcccaga 1140  
 tactttgtga agttcctgga tgcctatgct gagcacaagt tacagttctg ggcagtgaca 1200  
 gctgaaaatg agccttctgc tgggctgttg agtgatacc ccttccagtg cctgggcttc 1260  
 acccctgaac atcagcgaga cttcattgcc cgtgacctag gtccctacct cgccaacagt 1320  
 actcaccaca atgtccgcct actcatgctg gatgaccaac gcttgctgct gccccactgg 1380  
 gcaaagggtg tactgacaga cccagaagca gctaaatatg ttcattggcat tgctgtacat 1440  
 tggtagcttg actttctggc tccagccaaa gccaccctag gggagacaca ccgcctgttc 1500  
 cccaacacca tgctctttgc ctcagaggcc tgtgtgggct ccaagttctg ggagcagagt 1560

```

gtgctggctag gctcctggga tcgagggatg cagtacagcc acagcatcat cacgaacctc 1620
ctgtaccatg tggctggctg gaccgactgg aaccttgccc tgaaccccgaggaggaccc 1680
aattgggtgc gtaactttgt cgacagtccc atcattgtag acatcaccaa ggacacgttt 1740
tacaaacagc ccatgttcta ccaccttggc cacttcagca agttcattcc tgagggctcc 1800
cagagagtgg ggctgggtgc cagtcagaag aacgacctgg acgcagtggc actgatgcat 1860
cccgatggct ctgctgttgt ggtcgtgcta aaccgctcct ctaaggatgt gcctcttacc 1920
atcaaggatc ctgctgtggg cttcctggag acaatctcac ctggctactc cattcacacc 1980
tacctgtggc atcgccaaga tcttttagtc gatactatgt aatttcatga tctgttttgt 2040
tgtattccct tgcaatgcag ggcctagggc tatgaataaa gttaatgtgt gaatgtgtga 2100
atgtgtgatt gtgacctgaa gggatcacga ctataatcgt ttataataaa caaagacttt 2160
gtcccaaaaa ccccccccc ngcaga 2186

```

```

<210> 14
<211> 526
<212> PRT
<213> Artificial sequence

```

```

<220>
<223> High mannose human glucocerebrosidase (GCD)

```

```

<400> 14

```

```

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1           5           10           15

```

```

Leu Ser Ser Ala Glu Phe Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly
20           25           30

```

```

Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe
35           40           45

```

```

Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser
50           55           60

```

```

Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala
65           70           75           80

```

```

Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys
85           90           95

```

```

Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala Ala
100          105          110

```

```

Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys
115          120          125

```

```

Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro
130          135          140

```

```

Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr
145          150          155          160

```

Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp Thr  
 165 170 175  
 Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln Arg  
 180 185 190  
 Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys  
 195 200 205  
 Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly  
 210 215 220  
 Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp  
 225 230 235 240  
 Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu Asn  
 245 250 255  
 Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly  
 260 265 270  
 Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro  
 275 280 285  
 Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu Asp  
 290 295 300  
 Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr Asp  
 305 310 315 320  
 Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr Leu  
 325 330 335  
 Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg Leu  
 340 345 350  
 Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys  
 355 360 365  
 Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln  
 370 375 380  
 Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp  
 385 390 395 400  
 Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val  
 405 410 415  
 Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr  
 420 425 430  
 Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe



435

440

445

Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn  
 450 455 460

Asp Leu Asp Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val  
 465 470 475 480

Val Val Leu Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp  
 485 490 495

Pro Ala Val Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His  
 500 505 510

Thr Tyr Leu Trp His Arg Gln Asp Leu Leu Val Asp Thr Met  
 515 520 525